



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/864,486
Source: Q/PE
Date Processed by STIC: 12/5/03

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/864,486

DATE: 12/05/2003

TIME: 12:39:17

Input Set : A:\PTO.PG.txt

Output Set: N:\CRF4\12052003\I864486.raw

4 <110> APPLICANT: CENTER FOR GENETIC ENGINEERING AND BIOTHECNOLOGY
 6 <120> TITLE OF INVENTION: DNA fragments of the methylotrophic Pichia pastoris yeast

ICL

7 gene
 9 <130> FILE REFERENCE: List of Sequences (English)
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/09/864,486
 C--> 12 <141> CURRENT FILING DATE: 2001-05-24
 14 <160> NUMBER OF SEQ ID NOS: 10
 16 <170> SOFTWARE: PatentIn Ver. 2.1
 18 <210> SEQ ID NO: 1
 19 <211> LENGTH: 684
 20 <212> TYPE: DNA
 21 <213> ORGANISM: Pichia pastoris
 23 <220> FEATURE:
 24 <221> NAME/KEY: promoter
 25 <222> LOCATION: (1)..(684)
 26 <223> OTHER INFORMATION: Sequence that contains the 5' regulating region of ICL gene.
 28 <400> SEQUENCE: 1
 29 gaattcggac aaatgtgctg ttccggtagc ttgtaggaag cggcatccgt agggcaatat 60
 W--> 30 acgactatag cttctaaagc gtagtacaat gaaatgttcg aaggaaacaan aaacggattt 120
 31 gtttttcgta ggctcaacch gttgaggtgt aactcttttag ngaaagggtg agattgattg 180
 32 ttccgaagtag gggcctcaaa gggaaagaga aaaaaaaaaa tacaccgaag agttacgtaa 240
 33 gcatatattt ttacgtaaa gcatgattga atttcagcag tattgtttta caaggctgat 300
 34 gtcgtgtgcc aatcaaaaca aaagagattc gcataatgcc ataattgggg tgtgtgggg 360
 35 tccccnaaaa cgtctttctc atcatcatct gcaacccccca tcgaacctca ttaaatacaca 420
 36 tgacttgtgc gatcctcggt caactcggtc cgtgcaccca ttccaccccg ggctgaccaa 480
 37 cgcaagggtt tccgagagtc cgctacccca gatttatatc agcaaccagt cacttttttc 540
 38 cgggcacgac tctatatgcc ctggaaaacc ggagacgatg agcctgacta taaaagggtga 600
 39 cagaaccccc aactctggtt aatctcttca acaaatactt tttttcttt caattcaaag 660
 40 aacacagtat caagtatatc aaga 684
 43 <210> SEQ ID NO: 2
 44 <211> LENGTH: 360
 45 <212> TYPE: DNA
 46 <213> ORGANISM: Pichia pastoris
 W--> 47 <220> FEATURE:
 48 <221> NAME/KEY: terminator
 49 <222> LOCATION: (1)..(360)
 50 <223> OTHER INFORMATION: Sequence that contains the 3' regulating region of ICL gene.
 52 <400> SEQUENCE: 2
 53 gtaataggag ttcttaagta gtttaagataa ttgacttgag gtattttatag atttgtgtgt 60
 54 aggtaatatc tatggtcgtc cattcttacc ttgggtgggt gacggggcgg tgaataaatc 120
 55 agttgcgatg aaagacttta caaccttgtc accagagggt gcggtctact gattactaca 180
 56 aacgacttgg ataaaatttt caattcaaaa tcaatataaa aaaaaaaact taacatcact 240
 57 gatgtttcac taaactcttt aaacgctcaa ctcagcttc caactcgctc ttgcaaatca 300

pp 1-4
 Does Not Comply
 Corrected Diskette Needed

n's need to
 be explained
 (see p. 4
 for error
 explanation)

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/864,486

DATE: 12/05/2003

TIME: 12:39:17

Input Set : A:\PTO.PG.txt

Output Set: N:\CRF4\12052003\I864486.raw

58 gtaactcctc aactttgtct tccagttgac tcattctctt catcttctta gccctggaac 360

61 <210> SEQ ID NO: 3

62 <211> LENGTH: 21

63 <212> TYPE: DNA

64 <213> ORGANISM: Artificial Sequence

66 <220> FEATURE:

67 <223> OTHER INFORMATION: Description of Artificial Sequence:

68 oligonucleotide

70 <400> SEQUENCE: 3

71 tchggwtggc artgytcht h 21

74 <210> SEQ ID NO: 4

75 <211> LENGTH: 21

76 <212> TYPE: DNA

77 <213> ORGANISM: Artificial Sequence

79 <220> FEATURE:

80 <223> OTHER INFORMATION: Description of Artificial Sequence:

81 oligonucleotide

83 <400> SEQUENCE: 4

W--> 84 ~~gtncahgarg~~ aycarttyaa r *see p. 4* 21

87 <210> SEQ ID NO: 5

88 <211> LENGTH: 27

89 <212> TYPE: DNA

90 <213> ORGANISM: Artificial Sequence

92 <220> FEATURE:

93 <223> OTHER INFORMATION: Description of Artificial Sequence:

94 oligonucleotide

96 <400> SEQUENCE: 5

97 ctgcaggaat taattgcct tagacat 27

100 <210> SEQ ID NO: 6

101 <211> LENGTH: 26

102 <212> TYPE: DNA

103 <213> ORGANISM: Artificial Sequence

105 <220> FEATURE:

106 <223> OTHER INFORMATION: Description of Artificial Sequence:

107 oligonucleotide

109 <400> SEQUENCE: 6

110 aagcttgctg taacgaatct agaact 26

113 <210> SEQ ID NO: 7

114 <211> LENGTH: 23

115 <212> TYPE: DNA

116 <213> ORGANISM: Artificial Sequence

118 <220> FEATURE:

119 <223> OTHER INFORMATION: Description of Artificial Sequence:

120 oligonucleotide

122 <400> SEQUENCE: 7

123 cctcgagctt gtaggaattc gca 23

126 <210> SEQ ID NO: 8

127 <211> LENGTH: 24

128 <212> TYPE: DNA

RAW SEQUENCE LISTING

DATE: 12/05/2003

PATENT APPLICATION: US/09/864,486

TIME: 12:39:17

Input Set : A:\PTO.PG.txt

Output Set: N:\CRF4\12052003\I864486.raw

129 <213> ORGANISM: Artificial Sequence
131 <220> FEATURE:
132 <223> OTHER INFORMATION: Description of Artificial Sequence:
133 oligonucleotide
135 <400> SEQUENCE: 8
136 aaggtgctag cattcttgat atac 24
139 <210> SEQ ID NO: 9
140 <211> LENGTH: 24
141 <212> TYPE: DNA
142 <213> ORGANISM: Artificial Sequence
144 <220> FEATURE:
145 <223> OTHER INFORMATION: Description of Artificial Sequence:
146 oligonucleotide
148 <400> SEQUENCE: 9
149 atgctagcgc aagctttcct tttc 24
152 <210> SEQ ID NO: 10
153 <211> LENGTH: 24
154 <212> TYPE: DNA
155 <213> ORGANISM: Artificial Sequence
157 <220> FEATURE:
158 <223> OTHER INFORMATION: Description of Artificial Sequence:
159 oligonucleotide
161 <400> SEQUENCE: 10
162 agctcgcgat cagctaattct gccca 24
W--> 168 Lic. Argia Poveda Marcheco
W--> 169 Representante Legal, CIGB *delete*

VARIABLE LOCATION SUMMARY

DATE: 12/05/2003

PATENT APPLICATION: US/09/864,486

TIME: 12:39:18

Input Set : A:\PTO.PG.txt

Output Set: N:\CRF4\12052003\I864486.raw

Use of n's or Xaa's (NEW RULES): *error explanation*

Use of n's and/or Xaa's have been detected in the Sequence Listing.

Use of <220> to <223> is MANDATORY if n's or Xaa's are present.

in <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

Seq#:1; N Pos. 110,140,161,359,361,366

Seq#:4; N Pos. 3

VERIFICATION SUMMARY

PATENT APPLICATION: **US/09/864,486**

DATE: 12/05/2003

TIME: 12:39:18

Input Set : **A:\PTO.PG.txt**

Output Set: **N:\CRF4\12052003\I864486.raw**

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:30 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:60
M:341 Repeated in SeqNo=1
L:47 M:283 W: Missing Blank Line separator, <220> field identifier
L:84 M:258 W: Mandatory Feature missing, <221> Tag not found for SEQ ID#:4
L:84 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:4
L:84 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0
L:168 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:3
L:169 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:10
L:169 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:3